INTERNATIONAL STANDARD

ISO/IEC 11179-4

First edition
1995-07-01

Information technology — Specification and standardization of data elements —

Part 4:
Rules and guidelines for the formulation of data definitions

Technologies de l'information — Spécification et normalisation des données —

Partie 4: Règles et directives pour la formulation des définitions des données

Reference number
ISO/IEC 11179-4:1995(E)

Contents

ISO/IEC Copyright Office • Case postale 56 • CH-1211 Genève 20 • Switzerland

Printed in Switzerland

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC1. Draft international Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75% of the national bodies casting a vote.

International Standard ISO/IEC 11179-4 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, SC 14, *Data element principles*.

ISO/IEC 11179 consists of the following parts, under the general title *Information technology - Specification and standardization of data elements*:

- *Part 1: Framework for the generation and standardization of data elements*
- *Part 2: Classification of concepts for the identification of domains*
- *Part 3: Basic attributes of data elements*
- *Part 4: Rules and guidelines for the formulation of data definitions*
- *Part 5: Naming and identification principles for data elements*
- *Part 6: Registration of data elements*

Annex A of this part of ISO/IEC 11179 is for information only.

Introduction

The purpose of a data element definition is to define the meaning of a data element. To ensure consistency and quality, this part of ISO/IEC 11179 provides rules and guidelines for the construction of well-formed data element definitions.

Precise and unambiguous data element definitions are one of the most critical aspects of ensuring data shareability. When two or more parties exchange data, it is essential that all are in explicit agreement on the meaning of that data. One of the primary vehicles for carrying the data's meaning is the data element definition. Therefore, it is mandatory that every data element have a well-formed definition; one that is clearly understood by every user. Poorly formulated data element definitions foster misunderstandings and ambiguities and often inhibit successful communication.

This part of ISO/IEC 11179 contains both rules and guidelines. These apply to the formulation of definitions of data elements used in information processing systems and information interchange. Rules are mandatory and testable for compliance. Guidelines are principles that shall be followed. Objective test criteria can be established for rules. Conformance with guidelines is assessed by judgement of reasonableness.

The data element names in this part of ISO/IEC 11179 do not follow a particular syntax. The term *data element* refers to *data element type;* the shorter term is used for convenience.

Information technology - Specification and standardization of data elements

Part 4: Rules and guidelines for the formulation of data definitions

1 Scope

This part of ISO/IEC 11179 specifies rules and guidelines for constructing definitions for data elements. Only semantic structures of data element definitions are addressed; specifications for formatting the definitions are deemed unnecessary for the purposes of this part of ISO/IEC 11179.

Although these definitional rules and guidelines pertain to data elements, they can also be applied in formulating definitions for other types of data constructs such as entity types, entities, relationships, attributes, object types (or classes), objects, segments, composites, code entries, and messages.

The definitional rules and guidelines in this part of ISO/IEC 11179 do not always apply to terminological definitions found in glossaries and language dictionaries. Differences exist between the rules that apply in a language dictionary, and the rules that apply in a data dictionary. For example, terms in a language dictionary may have multiple definitions in the dictionary, whereas data dictionary definitions must be unique within a dictionary and have a single meaning.

Many data element definitions include terms that themselves need to be defined (e.g., "charge", "allowance", "delivery"). Some of these terms may have different definitions in different industrial sectors. Therefore, there is a need for most data dictionaries to establish an associated *glossary* of terms used in the definitions. The area(s) of use for each term is identified in the glossary.

2 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this part of ISO/IEC 11179. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this part of ISO/IEC 11179 are encouraged to apply the most recent editions of the standard indicated below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO 704:1987 [1], *Principles and methods of terminology.*

ISO 1087:1990 [1], *Terminology - Vocabulary.*

ISO 2382-4:1987, *Information processing systems - Vocabulary - Part 4: Organization of data.*

[1] Currently under revision.

ISO 9735:1988, *Electronic data interchange for administration, commerce and transport (EDIFACT) - Application level syntax rules.*

ISO 10241:1992, *International terminology standards - Preparation and layout.*

ISO/IEC 11179-3:1994, *Information technology - Specification and standardization of data elements - Part 3: Basic attributes of data elements.*

ANSI X3.172-1990, *American National Standard Dictionary for Information Systems (ANDIS).*

3 Definitions

For the purposes of this part of ISO/IEC 11179, the following definitions apply.

3.1 attribute: A characteristic of an object or entity.

3.2 concept: A unit of thought constituted through abstraction on the basis of characteristics common to a set of objects. This definition may apply to other types of data constructs as identified in the Scope of this part of ISO/IEC 11179.

3.3 data: A representation of facts, concepts, or instructions in a formalized manner suitable for communication, interpretation, or processing by humans or by automatic means (ISO 2382-4).

3.4 data dictionary: A database used for data that refers to the use and structure of other data; that is, a database for the storage of metadata (ANSI X3.172-1990). See *data element dictionary*.

3.5 data element: A unit of data for which the definition, identification, representation and permissible values are specified by means of a set of *attributes*.

3.6 data element dictionary: An information resource that lists and defines all relevant *data elements*. Synonymous with data dictionary.

NOTE - Data element dictionaries may exist at various levels, e.g. ISO/IEC Committees, international associations, industry sectors, companies, application systems.

3.7 definition: A word or phrase expressing the essential nature of a person or thing or class of persons or things: an answer to the question "what is x?" or "what is an x?"; a statement of the meaning of a word or word group [Webster's Third New International Dictionary of the English Language Unabridged, 1986].

3.8 domain: The set of possible data values of an attribute (ANSI X3.172-1990).

3.9 name: The primary means of identification of objects and concepts for humans.

4 Summary of data definition rules and guidelines

A listing of the rules and guidelines without explanations is provided in this clause for convenience of the user. The intent is to facilitate ease of use of this document once an understanding of the rules and guidelines is achieved. Clause 5 describes each rule and guideline with an explanation and examples to ensure their exact meaning is understood.

4.1 Rules

A data definition shall:

a) be unique (within any data dictionary in which it appears)
b) be stated in the singular
c) state what the concept is, not only what it is not
d) be stated as a descriptive phrase or sentence(s)
e) contain only commonly understood abbreviations
f) be expressed without embedding definitions of other data elements or underlying concepts

4.2 Guidelines

A data definition should:

a) state the essential meaning of the concept
b) be precise and unambiguous
c) be concise
d) be able to stand alone
e) be expressed without embedding rationale, functional usage, domain information, or procedural information
f) avoid circular reasoning
g) use the same terminology and consistent logical structure for related definitions

5 Requirements

5.1 Premises

Data elements exist and are used for specific purposes. Differences in use will require different operational manifestations of some rules and guidelines. For example, different levels of specificity for data element definitions are generally required in different contexts. Guideline 5.3.a) below, provides an example of this need for varying levels of specificity for different definitions. The implementation of Guideline 5.3.a), "state the essential meaning of the concept" is context dependent. The primary characteristics deemed necessary to convey the essential meaning of a particular definition will vary according to the level of generalization or specialization of the data element. The primary characteristics should include consideration of the relevance of any object class, property and qualifier associated with the concepts being analyzed. Primary and essential characteristics for defining concepts such as "airport" in the commercial air transportation industry might be specific, where a more general definition may be adequate in a different context. For a discussion of relationships between concepts in different contexts and how characteristics are used to differentiate concepts, see ISO 704, Clause 3. Definitions should be written to facilitate understanding by any user and by recipients of shared data.

5.2 Rules

To facilitate understanding of the rules for construction of well-formed data element definitions, explanations and examples are provided below. Each rule is followed by a short explanation of its meaning. Examples are given to support the explanations. In all cases, a good example is provided to exemplify the explanation. When deemed beneficial, a poor, but commonly used example is given to show how a definition should **NOT** be constructed. To further explain the differences between the good and poor examples, examples are followed by a statement of rationale behind them.

A data definition shall:

a) be unique (within any data dictionary in which it appears)

EXPLANATION - Each definition shall be distinguishable from every other definition (within the dictionary) to ensure the specificity is retained. One or several characteristics expressed in the definition must differentiate the concept to be defined from other concepts.

EXAMPLE -

1) good definitions:

"Goods Dispatch Date" - Date on which goods are dispatched by a given party.
"Goods Receipt Date" - Date on which goods are received by a given party.

2) poor definitions:

"Goods Dispatch Date" - Date on which goods are delivered.
"Goods Receipt Date" - Date on which goods are delivered.

REASON - The definition "Date on which the goods are delivered" cannot be used for both data elements "goods receipt date" and "goods dispatch date". Instead, each definition must be different.

b) be stated in the singular

EXPLANATION - The concept expressed by the data definition shall be expressed in the singular. (An exception is made if the concept itself is plural.)

EXAMPLE - "Article Number"

1) good definition:

A reference number that identifies an article.

2) poor definition:

Reference number identifying articles.

REASON - The poor definition uses the plural word "articles", which is ambiguous, since it could imply that an "article number" refers to more than one article.

c) state what the concept is, not only what it is not

EXPLANATION - When constructing definitions, the concept cannot be defined exclusively by stating what the concept is **not.**

EXAMPLE - "Freight Cost Amount"

1) good definition:

Cost amount incurred by a shipper in moving goods from one place to another.

2) poor definition:

Costs which are not related to packing, documentation, loading, unloading, and insurance.

REASON - The poor definition does not specify what is included in the meaning of the data.

d) be stated as a descriptive phrase or sentence(s) (in most languages)

EXPLANATION - A phrase is necessary (in most languages) to form a precise definition that includes the essential characteristics of the concept. Simply stating one or more synonym(s) is insufficient. Simply restating the words of the name in a different order is insufficient. If more than a descriptive phrase is needed, use complete, grammatically correct sentences.

EXAMPLE - "Agent Name"

1) good definition:

Name of party authorized to act on behalf of another party.

2) poor definition:

Representative.

REASON "Representative" is a near-synonym of the data element name, which is not adequate for a definition.

e) contain only commonly understood abbreviations

EXPLANATION - Understanding the meaning of an abbreviation, including acronyms and initialisms, is usually confined to a certain environment. In other environments the same abbreviation can cause misinterpretation or confusion. Therefore, to avoid ambiguity, full words, not abbreviations, shall be used in the definition.

Exceptions to this rule may be made if an abbreviation is commonly understood such as "i.e." and "e.g." or if an abbreviation is more readily understood than the full form of a complex term and has been adopted as a term in its own right such as "radar" standing for "radio detecting and ranging".

All acronyms must be expanded on the first occurrence.

EXAMPLE 1 - "Tide Height"

1) good definition:

The vertical distance from mean sea level (MSL) to a specific tide level.

2) poor definition:

The vertical distance from MSL to a specific tide level.

REASON- The poor definition is unclear because the acronym, MSL, is not commonly understood and some users may need to refer to other sources to determine what it represents. Without the full word, finding the term in a glossary may be difficult or impossible.

EXAMPLE 2 - "Unit of Density Measurement"

1) good definition:

The unit employed in measuring the concentration of matter in terms of mass per unit (m.p.u.) volume (e.g., pound per cubic foot; kilogram per cubic meter).

2) poor definition:

The unit employed in measuring the concentration of matter in terms of m.p.u. volume (e.g. pound per cubic foot, kilogram per cubic meter).

REASON - M.p.u. is not a common abbreviation and its meaning may not be understood by some users. The abbreviation should be expanded to full words.

f) be expressed without embedding definitions of other data elements or underlying concepts

EXPLANATION - The definition of a second data element or related concept should not appear in the definition proper of the primary data element. Definitions of terms should be provided in an associated glossary. If the second definition is necessary, it may be attached by a note at the end of the primary definition's main text or as a separate entry in the dictionary. Related definitions can be accessed through relational attributes (e.g., cross-reference).

EXAMPLE 1- "Sample Type Code"

1) good definition:

A code identifying the kind of sample.

2) poor definition:

> A code identifying the kind of sample collected. A sample is a small specimen taken for testing. It can be either an actual sample for testing, or a quality control surrogate sample. A quality control sample is a surrogate sample taken to verify results of actual samples.

REASON - The poor definition contains two extraneous definitions embedded in it. They are definitions of "sample" and of "quality control sample".

EXAMPLE 2 - "Issuing Bank Documentary Credit Number"

1) good definition:

> Reference number assigned by issuing bank to a documentary credit.

2) poor definition:

> Reference number assigned by issuing bank to a documentary credit. A documentary credit is a document in which a bank states that it has issued a documentary credit under which the beneficiary is to obtain payment, acceptance, or negotiation on compliance with certain terms and conditions and against presentation of stipulated documents and such drafts as may be specified.

REASON - The poor definition contains a concept definition, which should be included in a glossary.

5.3 Guidelines (Guiding principles)

A data definition should:

a) state the essential meaning of the concept

EXPLANATION - All primary characteristics of the concept represented should appear in the definition at the relevant level of specificity for the context. The inclusion of non-essential characteristics should be avoided. The level of detail necessary is dependent upon the needs of the system user and environment.

EXAMPLE 1 - "Consignment Loading Sequence Number" (Intended context: any form of transportation)

1) good definition:

> A number indicating the sequence in which consignments are loaded in a means of transport or piece of transport equipment.

2) poor definition:

A number indicating the sequence in which consignments are loaded in a truck.

REASON - In the intended context, consignments can be transported by various transportation modes, e.g., trucks, vessels or freight trains. Consignments are not limited to trucks for transport.

EXAMPLE 2 - "Invoice Amount"

1) good definition:

Total sum charged on an invoice.

2) poor definition:

The total sum of all chargeable items mentioned on an invoice, taking into account deductions on one hand, such as allowances and discounts, and additions on the other hand, such as charges for insurance, transport, handling, etc.

REASON - The poor definition includes extraneous material.

b) be precise and unambiguous

EXPLANATION - The **exact** meaning and interpretation of the defined concept should be apparent from the definition. A definition should be clear enough to allow only one possible interpretation.

EXAMPLE - "Shipment Receipt Date"

1) good definition:

Date on which a shipment is received by the receiving party.

2) poor definition:

Date on which a specific shipment is delivered.

REASON - The poor definition does not specify what determines a "delivery". "Delivery" could be understood as either the act of unloading a product at the intended destination or the point at which the intended customer actually obtains the product. It is possible that the intended customer never receives the product that has been unloaded at his site or the customer may receive the product days after it was unloaded at the site.

c) be concise

EXPLANATION - The definition should be brief and comprehensive. Extraneous qualifying phrases such as "for the purpose of this data dictionary," "terms to be described," shall be avoided.

EXAMPLE - "Character Set Name"

1) good definition:

> The name given to the set of phonetic or ideographic symbols in which data is encoded.

2) poor definition:

> The name given to the set of phonetic or ideographic symbols in which data is encoded, for the purpose of this data dictionary, or, as used elsewhere, the capability of systems hardware and software to process data encoded in one or more scripts.

REASON - In the poor definition, all the phrases after "...data is encoded" are extraneous qualifying phrases.

d) be able to stand alone

EXPLANATION - The meaning of the concept should be apparent from the definition. Additional explanations or references should not be necessary for understanding the meaning of the definition.

EXAMPLE - "School Location City Name"

1) good definition:

> Name of the city where a school is situated.

2) poor definition:

> See "school site".

REASON - The poor definition does not stand alone, it requires the aid of a second definition (school site) to understand the meaning of the first.

e) be expressed without embedding rationale, functional usage, domain information, or procedural information

EXPLANATION - Although they are often necessary, such statements do not belong in the definition proper because they contain information extraneous to the purpose of the definition. If deemed useful, such expressions may be placed in other data element attributes (ISO/IEC 11179-3).

1. The rationale for a given definition should not be included as part of the definition (e.g. if a data element uses miles instead of kilometers, the reason should not be indicated in the definition).
2. Functional usage such as: "this data element should not be used for ..." should not be included in the definition proper.
3. Remarks about procedural aspects. For example, "This data element is used in conjunction with data element 'xxx'", should not appear in the definition; instead use "Related data reference" and "Type of relationship" as specified in ISO/IEC 11179-3.

EXAMPLE - "Data Field Label"

1) good definition:

Identification of a field in an index, thesaurus, query, database, etc.

2) poor definition:

Identification of a field in an index, thesaurus, query, database, etc., which is provided for units of information such as abstracts, columns within tables.

REASON - The poor definition contains remarks about functional usage. This information starting with "which is provided for..." must be excluded from the definition and placed in another attribute, if it is necessary information.

f) avoid circular reasoning

EXPLANATION - Two definitions shall not be defined in terms of each other. A definition should not use another concept's definition as its definition. This results in a situation where a concept is defined with the aid of another concept that is, in turn, defined with the aid of the given concept.

EXAMPLE - two data elements with poor definitions:

1) "Employee ID Number" - Number assigned to an employee.
2) "Employee" - Person corresponding to the employee ID number.

REASON - Each definition refers to the other for its meaning. The meaning is not given in either definition.

g) use the same terminology and consistent logical structure for related definitions

EXPLANATION - A common terminology and syntax should be used for similar or associated definitions.

EXAMPLE - The example for rule 5.2.a) above also illustrates this idea. Both definitions pertain to related concepts and therefore have the same logical structure and similar terminology.

1) "Goods Dispatch Date" - Date on which goods were dispatched by a given party.
2) "Goods Receipt Date" - Date on which goods were received by a given party.

REASON - Using the same terminology and syntax facilitates understanding. Otherwise, users wonder whether some difference is implied by use of synonymous terms and variable syntax.

Annex A
(informative)

Bibliography

ISO Standards Handbook 10, *Data Processing - Vocabulary*, 1982

TRADE/WP.4/R.765/Add. 1, United Nations Economic and Social Council, 30 July 1991.

This page intentionally left blank

This page intentionally left blank

ICS 35.040

Descriptors: data processing, information interchange, data representation, data elements, specifications, definitions, rules (instructions).

Price based on 13 pages